Journal of Immunological Methods 475 (2019) 112520

Contents lists available at ScienceDirect

Journal of Immunological Methods

journal homepage: www.elsevier.com/locate/jim

Technical Note

Frequent issues and lessons learned from EuroFlow QA

Tomas Kalina[a,*], Nadezda Brdickova[a], Hana Glier[b], Paula Fernandez[b], Marieke Bitter[c], Juan Flores-Montero[d], Jacques J.M. van Dongen[c,**], Alberto Orfao[d]

[a] CLIP Cytometry, Department of Pediatric Hematology and Oncology, 2nd Faculty of Medicine, Charles University and University Hospital Motol, Prague, Czech Republic
[b] FACS/Stem cell Laboratory, Institute of Laboratory Medicine, Kantonsspital Aarau AG, Aarau, Switzerland
[c] Department of Immunohematology and Blood Transfusion (IHB), Leiden University Medical Center (LUMC), Leiden, The Netherlands
[d] Cancer Research Centre (IBMCC, USAL-CSIC), Institute for Biomedical Research of Salamanca (IBSAL), Department of Medicine and Cytometry Service (NUCLEUS Research Support Platform), University of Salamanca (USAL), Salamanca, Spain

ARTICLE INFO

Keywords:
Flow cytometry
Standardization
EuroFlow
Quality assessment scheme

ABSTRACT

EuroFlow Quality Assessment was designed to provide a feedback on the quality of the standardization effort in executing the EuroFlow protocols for sample preparation and instrument setup. It was first beta-tested by the members of the EuroFlow consortium internally (2010 – 2013) and opened to the external participants from 2015 onwards. The goal of participation in the EuroFlow QA is to evaluate whether the technical quality of the data generated by the laboratory is comparable to the data of the EuroFlow members and thus if a non-EuroFlow member participant can use the EuroFlow reference sample database for his own patient evaluation. Also it assesses whether data are sufficiently standardized for automated population gating and alarm notification. By spring 2018, a total 87 laboratories from 32 countries on five continents have registered for the EuroFlow QA program. We evaluated 163 results of 2015–2016 QA rounds, where we noted clear improvement in the score of first-time participants (median score of 91% correct) when they participated second time or later (median score of 94% correct, $p = 0{,}017$), which was comparable to EuroFlow member scores (median score of 97% correct). Among frequent mistakes, we found non-adherence to the EuroFlow protocols (improper reagent used), improper gating and some compensation issues. In summary, we show that EuroFlow QA has a positive impact on improvement of standardized data quality of non-member laboratories adhering to the EuroFlow standard operating procedures and reagent panels.

1. Introduction

The EuroFlow consortium has developed a series of standardized protocols for diagnostic use (van Dongen et al., 2012) and for monitoring treatment response (Theunissen et al., 2016; Flores-Montero et al., 2017a) by flow cytometry. These protocols have been adopted in multiple diagnostic laboratories. 87 laboratories from 32 countries on five continents have registered for the EuroFlow QA program, which means they are actively using the EuroFlow protocols and trying to improve. The interest in the protocols is also reflected by the fact that around 16,000 documents have been downloaded by ~1750 institutes in 3 years-time from the euroflow.org website. Patients' samples acquired in this standardized fashion can be directly compared to each other, and carefully annotated measurements can be compiled in a data base (Pedreira et al., 2013; Pedreira et al., 2008) and used as a reference for evaluation of new cases (Lhermitte et al., 2018; Costa et al., 2010). Overall, 3124 reference samples acquired in 15 EuroFlow laboratories were uploaded by January 2017 to the EuroFlow server to be used for the construction of the EuroFlow databases. Standardization of the instrument set-up and panels of reagents, together with a uniform approach for data analysis is instrumental to make comparisons of individual samples to the common database (Kalina et al., 2012). Quality of the measured data is evaluated with EuroFlow Quality Assessment scheme (Kalina et al., 2015). In this approach, obtaining virtually identical or highly similar median fluorescence intensity (MFI) values for a particular cell subset labeled with a given fluorescent reagent is considered a standardization endpoint. The EuroFlow Quality Assessment (QA) program was launched in 2010 for the members of the EuroFlow consortium to ensure their adherence to the agreed standardization. In 2014, the results of the first three years of the QA

* Correspondence to: T. Kalina, Department of Pediatric Hematology and Oncology, 2nd Faculty of Medicine, Charles University, Prague, V Uvalu 84, 150 06 Prague 5, Czech Republic.
** Correspondence to: J.J.M. van Dongen, Dept. of IHB Room L3-50, Leiden University Medical Center, Albinusdreef 2, 2333, ZA, Leiden, The Netherlands.
E-mail addresses: tomas.kalina@lfmotol.cuni.cz (T. Kalina), J.J.M.van_Dongen@lumc.nl (J.J.M. van Dongen).

https://doi.org/10.1016/j.jim.2018.09.008
Received 4 October 2017; Received in revised form 3 April 2018; Accepted 14 September 2018
Available online 17 September 2018

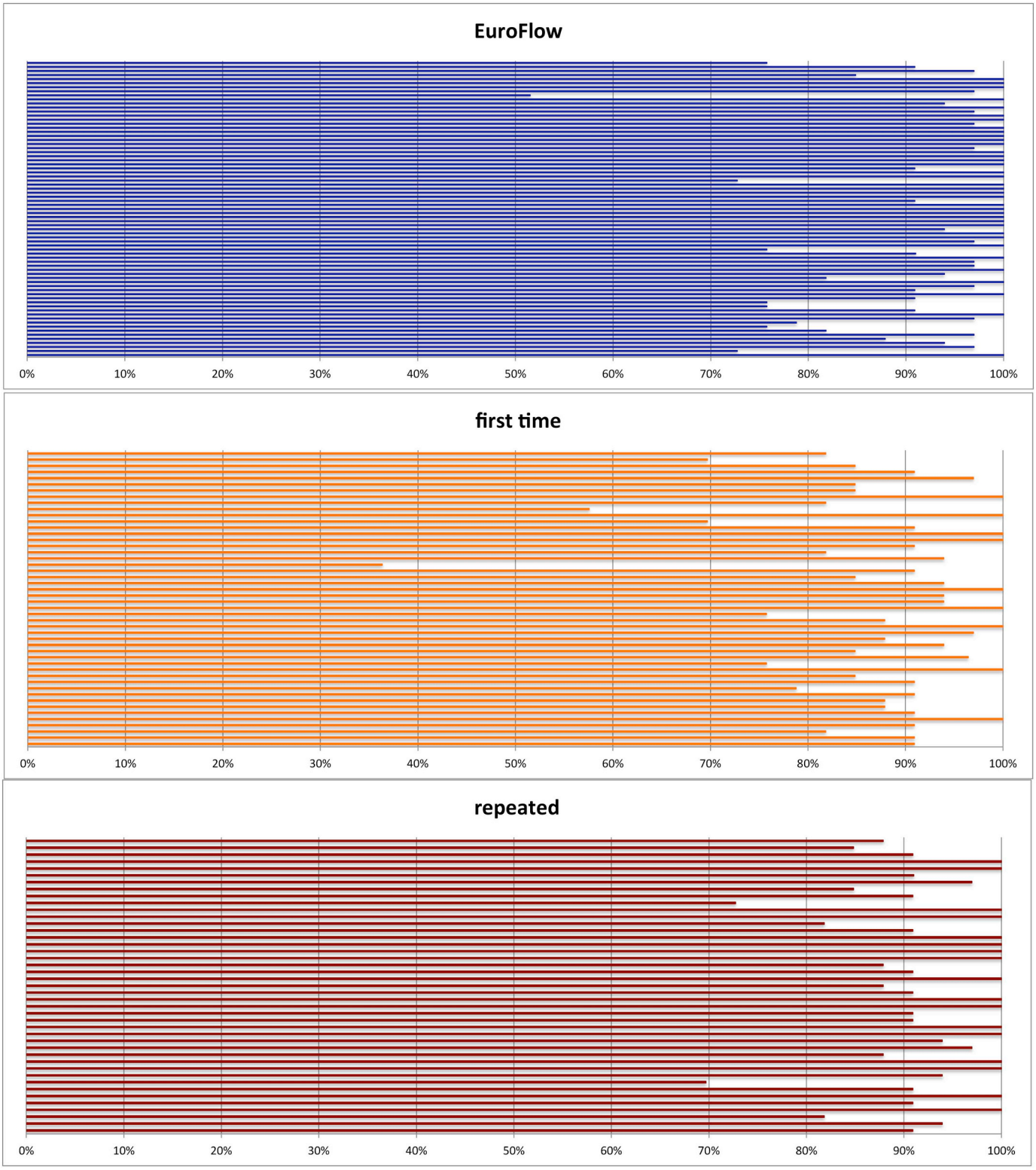

Fig. 1. Score of EuroFlow QA in 2015–2016 for A) EuroFlow members (blue), B) first participation of non-EuroFlow members (orange) and C) repeated participations of non-EuroFlow members (brown). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

program were summarized (Kalina et al., 2015) and in 2015 the EuroFlow QA was opened to participants outside the EuroFlow consortium. The EuroFlow QA was build to specifically test the quality of the execution of the QA modified Lymphoid Screening Tube (LST-QA) measurement, where locally drawn healthy volunteers' peripheral blood is used as a model. The guidance for the participants as well as reporting of the results is organized via a dedicated EuroFlow QA website and QA rounds are organized twice per year. The participants choose to use either the originally published single reagents or they can use a centrally distributed lyophilized mixture of reagents (both variants were shown to yield comparable results (van der Velden et al., 2017) and participants should choose the approach, which is more

Table 1
Frequency of tests failing per reagent included in the LST-QA combination.

Reagent	Out of range results
(reference cell population)	(n = 504)
CD20 PacB (B-cells)	3,8%
CD4 PacB (CD4 T-cells)	12,9%
CD45 PacO/OC515 (T-cells)	9,9%
CD8 FITC (CD8 T-cells)	3,0%
IgL FITC (B-cells)	5,0%
CD56 PE (NK-cells)	16,1%
IgK PE (B-cells)	9,7%
CD5 PerCP5–5 (T-cells)	4,4%
CD19 PECy7 (B-cells)	2,0%
CD3 APC (T-cells)	6,7%
CD81 APCH7/APCC750 (B-cells)	20,2%

resembling their routine practice) (Kalina et al., 2015). The evaluation of the QA-test is based on the distance of the measured MFI of predefined lymphocyte subsets labeled with the LST-QA reagents from the expected MFI established during the original standardization process. In contrast, the EuroFlow QA does not test the ability of the laboratory to perform a clinical interpretation (establish a diagnostic conclusion). Participants receive QA certificates with an overview of their results for all tested parameters. Whenever found, values out of range are highlighted. Participants also receive a graphical representation of the results highlighting the values measured by their laboratory among all laboratories participating in a given round. A summary of the issues encountered and their possible causes is circulated to all participants at the end of each round. In addition, the EuroFlow QA participants' meetings (first organized in Aarau, Switzerland, in November 2016) serve as a platform for education and troubleshooting.

Here we report on the EuroFlow QA program results and both the most frequent problems encountered among the participants in the EuroFlow QA program and the lessons learned from them.

2. Methods

2.1. Data acquisition, participants and instruments

Detailed execution of the EuroFlow QA has been described earlier (Kalina et al., 2015). The EuroFlow QA scheme was opened to non-EuroFlow member laboratories in 2015. By the end of 2016, a total 65 laboratories from 30 different countries in five continents have registered for the QA program (Supplementary Table 1) and 64 have been evaluated; the remaining laboratory did not submit QA results. Fifteen of them were EuroFlow member laboratories and fifty were non-members. Thus, out of 91 reports submitted by non-members, 46 were from first time participants, 45 were from repeated QA. 72 reports were from EuroFlow laboratories (three laboratories reported two sets of measurements for two different instruments each). In total 489 files have been evaluated (163 QA sets of three samples each). The majority of users acquired their data on the BD FACS Canto II (BD Biosciences, San Jose, CA) instruments, while two participating laboratories used the Navios instrument (Beckman Coulter, Brea, CA).

3. Data analysis

Analysis of each file was done according to the QA instructions by each reporting analyst and MFI values were entered into the EuroFlow QA website interface; in parallel the gated flow cytometry standard (FCS) data files analyzed with the Infinicyt software (Cytognos SL, Salamanca, Spain) ".cyt" were uploaded to the EuroFlow QA server. Meta-analysis of all reports and its graphical representations were done in Microsoft Excel for Mac 2011 (Microsoft Corporation, Redmond, Washington, USA).

4. Results

A comparison of the overall scores of the EuroFlow members and non- EuroFlow members is shown in Fig. 1. Overall score is calculated as a sum of p-score points falling into acceptable range of Dmax. P-score evaluates a deviation from expected (ideal) Median Fluorescence Intensity and is derived from the 90th (Kappa and Lambda) and 95th percentiles (other markers), thus it is expected that 90% of p-score values should fall within the acceptable range. The results of the experienced EuroFlow laboratories were close to perfect, with median score 97% (76%–100%; 10th–90th percentile). When only the first-time participants (non-EuroFlow members) were evaluated, they scored at median overall score of 91% (76%–100%). An improvement (p = .017) was achieved in the repeated reports (second and later QA rounds), where the non-members scored close to the EuroFlow members with median score of 94% (85%–100%). Of note, while the majority of users acquired their data on the BD FACSCanto II instruments, two participating laboratories used the Navios instrument successfully, confirming that the standardized approach can be taken across instruments of different brands without compromising the quality as described by Kalina et al. in this issue (Nováková et al., 2017), whenever the EuroFlow standard operating procedures (SOPs) and the EuroFlow recommended reference (or alternative) reagents are used.

For each parameter, we also analyzed the frequency of failure (Table 1). The most frequently failed parameter was CD81 APC-C750 (20.2%), followed by CD56 PE on CD56 bright NK cells (16.1%) and CD4 Pacific Blue (12.9%). Most CD81 failures were due to protocol non-adherence, where participants stained the sample with CD38 APCC750 instead (this is the only reagent different between the EuroFlow LST tube and the EuroFlow LST-QA tube). In turn, failure for the CD56 PE parameter on CD56-bright NK cells was typically due to improper gating. Since 82% of results were submitted together with their gated files for QA evaluation (recommended but not required), we could backtrack that instead of gating the CD56-bright cells they gated all CD56-positive cells. In the remaining cases we were not able to track down the reasons for the QA failures.

An example of how well the LST-QA tube with a perfect score matches a reference file (measured by different personnel, at a different site, on a different instrument, using different healthy adult normal peripheral blood samples) is shown in Fig. 2A. This is compared to a file with a score of 70% overall score in Fig. 2B.

5. Discussion

The EuroFlow QA program was successfully launched to public and it was well adopted by the laboratory diagnostics community. The same QA approach (using a complete EuroFlow LST tube) was adopted by the Swiss Cytometry Society for inter-laboratory training and evaluation (see Glier et al. in this issue (Glier et al., 2017)). Our experience so far shows that the EuroFlow QA program (Kalina et al., 2015) is properly set up, testing relevant endpoints and providing a feedback on the most frequent issues with the execution of the EuroFlow protocols. The consistency of signal intensity enables to discern the complex patterns, which form the basis for the comparison of individual samples to the EuroFlow database of known well-defined reference samples (Flores-

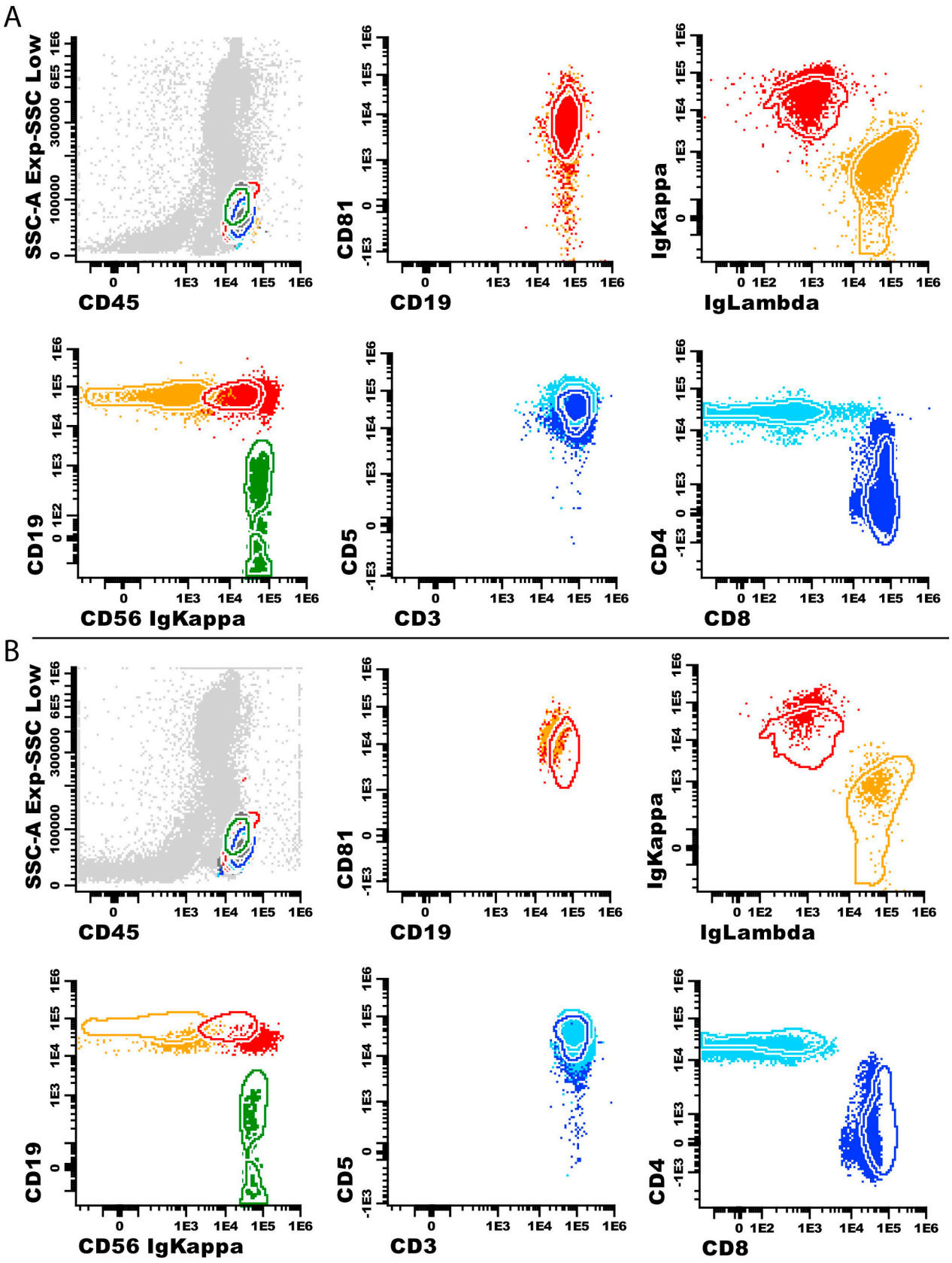

Fig. 2. Example of a perfectly correct result (Panel A) and a result with a score of 70% (Panel B). Results are shown as dots and expected patterns are depicted as reference image contours. Of note in panel B, CD4, CD8, CD19 and CD5 have lower intensity than expected).

Montero et al. in this issue (Lhermitte et al., 2018; Flores-Montero et al., 2017b)).

An essential component of the standardization of fluorescence intensity measurements is the reproducibility of the performance of antibody conjugates (Böttcher et al., 2017; Blanco et al., 2017). EuroFlow has built a list of reagents that perform equally well as the original reagents in a series of standardized measurements. Those are the so-called "EuroFlow alternative reagents" and their list is available at the www.euroflow.org website. To be included in the list, the alternative reagents had to show a similar staining pattern (in an overlay with the original reagent) and a median fluorescence intensity that does not differ by more than 30% from the original reagent when used on the same sample as a single reagent and/or when tested in the respective EuroFlow antibody combination in liquid or dried format (van der Velden et al., 2017). Testing was performed independently by three EuroFlow laboratories (in parallel) as a paid service to the reagents' manufacturers, and only valid alternative reagents being disclosed. While our data show that the alternative reagents can indeed perform equally well (at the time of testing), it is out of the control of EuroFlow or the end-user, whether a particular manufacturer maintains the reproducibility (to provide the same MFI in a standardized test) in different reagents' lots (Böttcher et al., 2017). Manufacturers often declare a Quality Control (QC) was performed, however there are no publically available guidelines or parameters that would allow the user to judge, whether the expected reproducibility of signal was achieved (and how it was measured). In this issue, Böttcher et al. (Böttcher et al., 2017), report on a lot-to-lot signal intensity evaluation using antibody capture beads, where the failure to provide a signal comparable to the previous lot ranged from 0% to 37.5%, for different fluorochromes and manufacturers. Thus, it is essential for standardized clinical investigations that the antibody producers perform a lot-to-lot intensity evaluation of their reagents and quantify and declare the reproducibility that is guaranteed by their QC process. In turn, it is implicitly stated here that, a given reference MFI value and not a staining intensity brighter than a given threshold MFI value is to be used as the criteria to define alternative reagents to the EuroFlow reference reagents.

Additional sources of differences we have observed were related to an improper execution of the sample preparation SOP (insufficient washing off the plasma) that led to a low labeling of B cells by antibodies against immunoglobulin (Ig) light chains (anti-lambda and anti-kappa Ig light chain reagents). In those samples, B cells had to compete for the labeling reagent with the abundant immunoglobulins in the unwashed plasma. Frequently, we noted issues that are caused by non-adherence to the protocol. A striking example is the usage of reagents that have never been tested and validated by the EuroFlow consortium and consequently, that have never been approved neither as reference nor as alternative reagents, and which do not perform similar to the reference EuroFlow reagents. On one hand, the degree of non-adherence is worrisome in clinical laboratories but at the same time this gets often corrected after receiving the feedback from the QA and retraining. As seen in other studies, we also point to improper gating strategies as a relatively frequent source of error (Finak et al., 2016; Whitby et al., 2012).

An imperfect compensation was another frequently found error. While the QA is not designed to formally test compensation issues, gross abnormalities can be easily recognized. Since compensation values are predictable in systems with fixed optical filters, standardized instrument setup and standardized panels of stable reagents, we expect that software solutions will be developed to reduce major compensation errors. Before this is achieved, a good practice in performing compensation experiments must be ensured through proper training.

Notably, because the QA program is based on a relatively stable expression of the evaluated surface markers, with very low variation among healthy individuals, there is no need to send out a reference sample. This not only makes it affordable at distant locations, but also logistically feasible across the globe, even in resource poor countries or countries with difficult custom procedures.

It should be stressed again that while the EuroFlow QA program tests the sample preparation, cytometer setting, aquisition and analysis, it does not test the proper diagnostic interpretation. Thus, it cannot replace the established QA programs required for ISO 15189 accreditation. However, this QA process together with the EuroFlow database and automated data analysis will be a significant aid to achieve also proper, robust and reproducible diagnoses.

Importantly, the EuroFlow QA involves a face-to-face participant meeting during which those issues found are reviewed with the participants, and offers retraining and education. The first educational meeting was organized in Aarau, Switzerland (November 2016), followed by Leiden (November 2017) and it will be repeated annually.

In conclusion, the EuroFlow QA program is well established and has proven its value for improving the execution of the EuroFlow protocols. Half of the non-EuroFlow laboratories scored well the first time they participated in the program and a further improvement could be clearly seen in subsequent rounds reaching scores comparable to the EuroFlow member laboratories.

Acknowledgements

All members of the EuroFlow consortium are acknowledged for their contribution to the network and fruitful discussions. This study was performed within the EuroFlow Consortium, the EU-supported LSHB-CT- 2006-018708 FP6 Specific Targeted Research Project, which obtained sustainability based on income from the development of intellectual property and related patents, which have been licensed to industry. The work of TK and NB was supported by grant 15-28525A from Ministry of Health of the Czech Republic and Instruments and infrastructure were supported by Ministry of Education, Youth and Sports, Czech Republic NPU I no. LO1604.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jim.2018.09.008.

References

Blanco, E., et al., 2017. Selection and validation of antibody clones against IgG and IgA subclasses in switched memory B-cells and plasma cells. J. Immunol. Methods.
Böttcher, S., et al., 2017. Lot-to-lot stability of antibody reagents for flow cytometry. J. Immunol. Methods. https://doi.org/10.1016/j.jim.2017.03.018.
Costa, E.S., et al., 2010. Automated pattern-guided principal component analysis vs expert-based immunophenotypic classification of B-cell chronic lymphoproliferative disorders: a step forward in the standardization of clinical immunophenotyping. Leukemia 24, 1927–1933.
Finak, G., et al., 2016. Standardizing Flow Cytometry Immunophenotyping Analysis from the Human ImmunoPhenotyping Consortium. Sci. Rep. 6, 20686.
Flores-Montero, J., et al., 2017a. Next Generation Flow for highly sensitive and standardized detection of minimal residual disease in multiple myeloma. Leukemia. https://doi.org/10.1038/leu.2017.29.
Flores-Montero, J., et al., 2017b. Data bases of normal samples as internal reference for automated gating and identification of abnormal cells. J. Immunol. Methods.
Glier, H., et al., 2017. Standardization of 8-color flow cytometry across different flow cytometer instruments: A feasibility study in clinical laboratories in Switzerland. J. Immunol. Methods. https://doi.org/10.1016/j.jim.2017.07.013.
Kalina, T., et al., 2012. EuroFlow standardization of flow cytometer instrument settings and immunophenotyping protocols. Leukemia 26, 1986–2010.
Kalina, T., et al., 2015. Quality assessment program for EuroFlow protocols: Summary results of four-year (2010 − 2013) quality assurance rounds. Cytom. Part A 87, 145–156.
Lhermitte, L., et al., 2018. Automated database-guided expert-supervised orientation for immunophenotypic diagnosis and classification of acute leukemia. Leukemia 32, 874–881.
Nováková, M., et al., 2017. How to make usage of the standardized EuroFlow 8-color protocols possible for instruments of different manufacturers. J. Immunol. Methods.
Pedreira, C.E., et al., 2008. A probabilistic approach for the evaluation of minimal residual disease by multiparameter flow cytometry in leukemic B-cell chronic lymphoproliferative disorders. Cytometry A 73A, 1141–1150.
Pedreira, C.E., Costa, E.S., Lecrevisse, Q., van Dongen, J.J.M., Orfao, A., 2013. Overview of clinical flow cytometry data analysis: recent advances and future challenges.

Trends Biotechnol. 31, 415–425.
Theunissen, P., et al., 2016. Standardized flow cytometry for highly sensitive MRD measurements in B-cell acute lymphoblastic leukemia. Blood. https://doi.org/10.1182/blood-2016-07-726307.
van der Velden, V.H.J., et al., 2017. Optimization and testing of dried antibody tube: The EuroFlow LST and PIDOT tubes as examples. J. Immunol. Methods. https://doi.org/10.1016/j.jim.2017.03.011.
van Dongen, J.J.M., et al., 2012. EuroFlow antibody panels for standardized n-dimensional flow cytometric immunophenotyping of normal, reactive and malignant leukocytes. Leukemia 26, 1908–1975.
Whitby, A., et al., 2012. ISHAGE protocol: Are we doing it correctly? Cytom. Part B Clin. Cytom. 82B, 9–17.